Stochastic simulation of epidemic spreading in scale-free networks

Luis Martin

March 2018

1 Introduction

Biological processes have an element of uncertainty attached to them. For example, humans reproduce some time between 15 and 40 years, but the exact time of reproduction for a given individual cannot be predicted. This randomness can be incorporated in the mathematical formalism. This project deals with the simulation and analysis of stochastic models in the context of epidemics.

Randomness is incorporated into the mathematical model as the probabilities at which an event occurs, and not, as in deterministic models, as rates. In the stochastic SIR model, this events will be the ones related with the epidemic processes, such as infection or death due to infection, but also every other process governing the demography. [1]

Epidemiology of directly transmitted infectious disease is fundamentally linked with networks. Early epidemiological models were based on population wide random-mixed, but in practice each individual has a finite set of contacts to whom they can pass infection; the ensemble of all such contacts forms a 'mixing network'. Knowledge of the structure of the network allows models to compute the epidemic dynamics at the population scale from the individual-level behaviour of infections. Therefore, we will also observe the main characteristics of mixing networks, and how these deviate from the random-mixing norm. [2]

2 The SIR model

The model consists of three compartments: S for the number susceptible, I for the number of infectious, and R for the number recovered (or immune). During an epidemic, the number of susceptible individuals falls rapidly as more of them are infected and thus enter the infectious and recovered compartments. [3] The model with which we are working is:

$$S + I \rightarrow 2I \tag{1}$$

$$I \rightarrow R \tag{2}$$

For the full specification of the model, the arrows should be labeled with the transition rates between compartments. Between S and I, the transition rate is βI, where β is the contact rate, which takes into account the probability of getting the disease in a contact between a susceptible and an infectious subject. Between I and R, the transition rate is γ (simply the rate of recovery or death). If the duration of the infection is denoted D, then $\gamma = \frac{1}{D}$, since an individual experiences one recovery in D units of time. For all this project $\gamma = \frac{1}{10} days^{-1}$.

The dynamical system can be interpreted by the following deterministic differential equations:

$$\frac{dS}{dt} = -\frac{\beta IS}{N} \tag{3}$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I \tag{4}$$

$$\frac{dR}{dt} = \gamma I \tag{5}$$

Firstly note that from $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$, it follows that $S(t)+I(t)+R(t) = N$, expressing in mathematical terms the constancy of population N. Note that this relationship implies that one need only study the equation for two of the three variables. Secondly note that the dynamics of the infectious class depends on the following ratio:

$$R_0 = \frac{\beta}{\gamma}$$

the so-called basic reproduction number (also called basic reproduction ratio). This ratio is derived as the expected number of new infections (these new infections are sometimes called secondary infections) from a single infection in a population where all subjects are susceptible.

It yields that if $R_0 > \frac{N}{S(0)}$, then $\frac{dI(0)}{dt} > 0$, i.e., there will be a proper epidemic outbreak with an increase of the number of the infectious (which can reach a considerable fraction of the population). On the contrary, if $R_0 < \frac{N}{S(0)}$, then $\frac{dI(0)}{dt} < 0$, i.e., independently from the initial size of the susceptible population the disease can never cause a proper epidemic outbreak. As a consequence, it is clear that the basic reproduction number is extremely important.

3 The stochastic SIR model

In the stochastic version of the SIR model, the continuous variables are replaced by discrete numbers, and the process rates are replaced by process probabilities. The master equation for the infected population given by:

$$\partial_t p(n,t) = (E-1)\gamma n p(n,t) + (E^{-1}-1)\frac{\beta}{V} S n p(n,t) \tag{6}$$

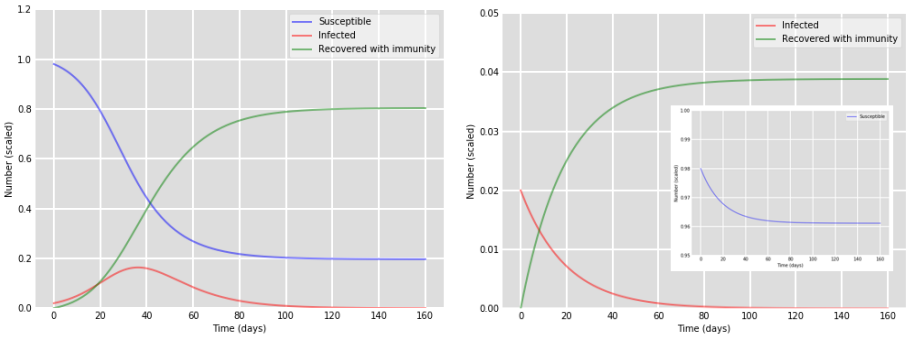

Figure 1: Integration of the SIR model equations. **Left:** The basic reproduction number ($\beta = 0.2 days^{-1}$) leads to an epidemic outbreak with an increase in the number of infected population, who reaches his maximum at the day 40. **Right:** The basic reproduction number ($\beta = 0.05 days^{-1}$) is not large enough to generate an epidemic outbreak. The infected population reaches its maximum at the beginning.

where V is the volume of he system where the dynamics takes place and S the number of susceptibles. The presence of the volume is related with the probability of an infected and a susceptible to meet in a mean field approximation. [4]

Although conceptually our stochastic SIR model is more difficult than the deterministic one, it is not more difficult to simulate. This is due to the fact that the discrete variables are easier for computers to handle than continuous ones.

The Gillespie algorithm is a very efficient, but still accurate way to simulate stochastic models. The idea of the Gillespie algorithm is that one first determines when something happens next. Suppose the current time is t. The time $t + \tau$ at which something happens next is an exponentially distributed random number scaled by the sum of all process rates.

$$\tau = \frac{-\ln u_0}{\sum_{j \neq i} w_{i \to j}} \tag{7}$$

where u_0 is an uniform random variable $\hat{\mathbf{U}}(0, 1)$. In our case:

$$\tau = \frac{-\ln u_0}{\frac{\beta}{V} Sn + \gamma n} \tag{8}$$

Then, the Gillespie algorithm determines what happens next. This is done by drawing a process randomly from all possible processes according to their respective probabilities. When we have determined which process happens, we can update the variables (the so-called state of the system). Then we iterate this process as long as we want. [4]

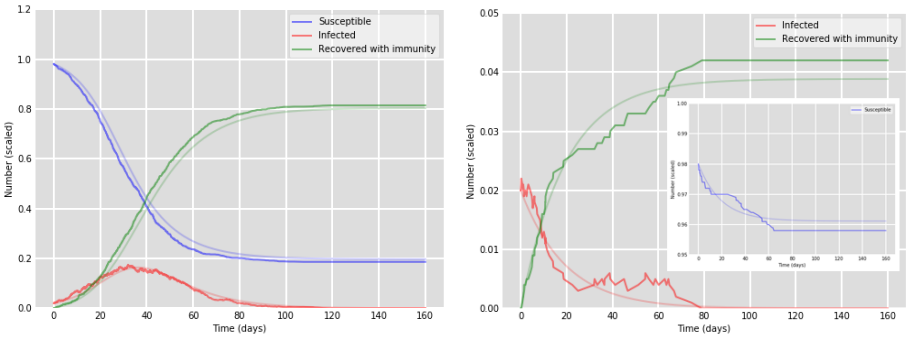

Figure 2: Integration of the stochastic SIR model using a Gillespie algorithm. The model has 10^3 individuals. (Soft lines are the mean field solution). **Left:** The basic reproduction number ($\beta = 0.2 days^{-1}$) leads to an epidemic outbreak with an increase in the number of infected population, who reaches his maximum at the day 40. **Right:** The basic reproduction number ($\beta = 0.05 days^{-1}$) is not large enough to generate an epidemic outbreak. The infected population reaches its maximum at the beginning.

In Fig.2 we have tried to reproduce the data obtained in Fig.1 solving the stochastic SIR model with a Gillespie algorithm. They are pretty accurate, but what we can deduce is that both models differ as long as the size of our system decreases. For a low infected population the deterministic equations fail to reproduce the real behaviour of the model. While in the stochastic model we take into account each individual, in the deterministic what we get is the mean evolution, which only corresponds to the real solution when the number of individuals in the system is large and the importance of the stochastic fluctuations is reduced. On the other hand, averaging over a set of solutions of the stochastic model we would reach something similar to the deterministic evolution. We could understand the solutions of the master equation as a deterministic component given by the differential equations of the model plus stochastic fluctuations, which lose importance as long as the size of the system increases.

4 The SIR model in a Network

In a real epidemic spreading process, the susceptible individuals will be infected only if they have contact with some infected ones. Until now we have modeled this contact using a mean field approach, both in the deterministic and in the stochastic models. What it is proposed next is to study how the infection would spread in a static network, where individuals are represented by nodes and they are infected with probability β for every infected neighbor they have. In this way, the classical epidemic models can be natural extended to network epidemic models.

We will use again the Gillespie algorithm to solve the master equation of the network SIR model, but now the probabilities of changing from one state to another will be individual dependent. We will need to have a rigorous control of the sate of every node and the state of his neighbors in order to compute τ and the transition rates. [5]

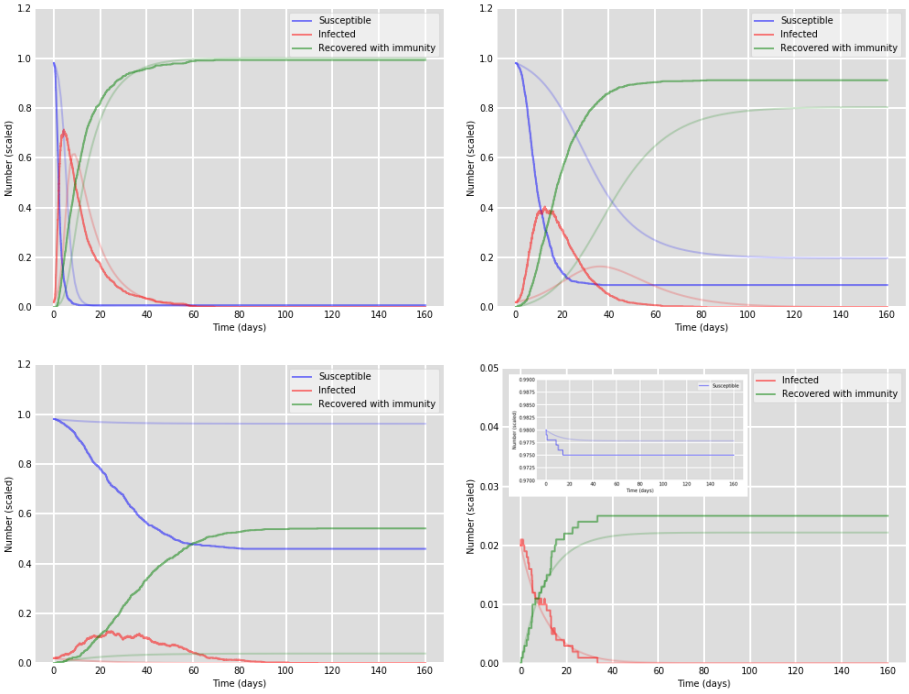

Figure 3: Integration of the network SIR model using a Gillespie algorithm in a *Barabassi-Albert* network of 10^3 nodes. (Soft lines are the mean field solution). **A:** With $\beta = 0.8$, R_o leads to an epidemic outbreak with an increase in the number of infected population, who reaches his maximum at the day 5. **B:** For $\beta = 0.2$, R_o stills above the threshold, but with less infectivity. The epidemic peak is reached at the day 15. **C:** R_o is below the threshold for $\beta = 0.05$, but we still having epidemic outbreak.**D:** R_o is below the threshold for $\beta = 0.01$, and we don't have epidemic outbreak.

In Fig.3 we can observe how the model behaves for different values of β in a scale-free network. It is known that social networks can be modeled using this kind of graphs. Even the dynamics of malware can be modeled using epidemic spreading in the Internet, which can be also modeled as a scale-free network. [6] The first thing that we can observe, in Fig.3**C** and Fig.3**D** is that we still having the epidemic threshold, but it can't be defined as before, since we still having

epidemic outbreak for $\beta = 0.05$. The threshold value is network dependent. On the other hand, the infectivity still defining where the outbreak takes place. When β is small the epidemic peak takes more time to arrive and also with less intensity. But in comparison with the deterministic results the spreading of the infection in the network is faster and reaches higher peak values. This is due to the topology of the scale-free network, which favours the spreading more than the mean field approach due to the high clustering coefficient.

5 The demographic SIR model

The most common modification of the SIR model previous analyzed is the SIR model with vital dynamics. [7] Now we consider a population characterized by a death rate μ and birth rate Λ, and where a communicable disease is spreading. The model with mass-action transmission is:

$$\frac{dS}{dt} = \Lambda N - \mu S - \frac{\beta IS}{N} \tag{9}$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - (\gamma + \mu)I \tag{10}$$

$$\frac{dR}{dt} = \gamma I - \mu R \tag{11}$$

In this case, we can derive a basic reproduction number :

$$R_0 = \frac{\beta \Lambda}{\mu(\mu + \gamma)} \tag{12}$$

which has threshold properties. In fact, independently from biologically meaningful initial values, one can show that $R_0 \leq 1$ leads to an equilibrium where there are no infected and removed, and the susceptible population is the ratio between newborns and deaths $S(t) = \frac{N\Delta}{\mu}$. If this ratio is one then the population N is constant. On the other hand, for $R_0 > 1$ we achieve the endemic equilibrium given by:

$$(S(t), I(t), R(t)) = \left(N\frac{\gamma + \mu}{\beta}, N\frac{\mu}{\beta}\left(R_0 - 1\right), N\frac{\gamma}{\beta}\left(R_0 - 1\right) \right) \tag{13}$$

With heuristic arguments, one may show that R_0 may be read as the average number of infections caused by a single infectious subject in a wholly susceptible population, the above relationship biologically means that if this number is less or equal than one the disease goes extinct, whereas if this number is greater than one the disease will remain permanently endemic in the population. In this work we will work with a constant population N doing $\Lambda = \mu = 1/50 days^{-1}$.

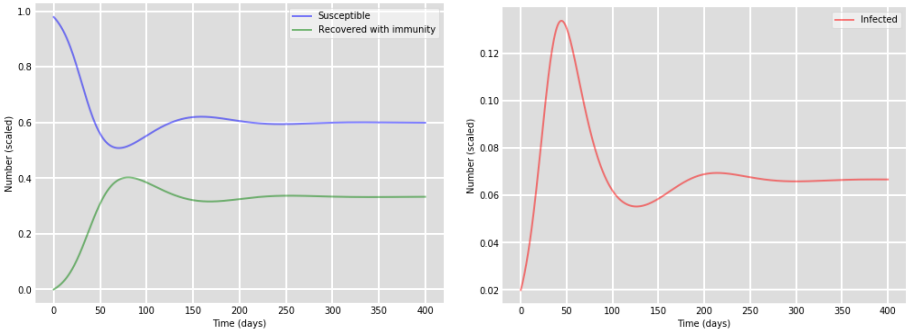

Figure 4: Integration of the demographic SIR model differential equations. For $\beta = 0.2$ the system reaches, oscillating, the endemic equilibrium at the day 400. It has an epidemic outbreak close to the day 50. The population is maintained constant using $\mu = \Delta$.

In Fig.5 we solve the stochastic system using the Gillespie. In this case we can observe more clearly that as long as the size of the system increase (the number of individuals in our simulations), the stochastic solutions of the master equation are more similar to the solution obtained from the system of differential equations. In this sense, Fig.5**B:** is much more similar to the deterministic solution plotted in Fig.4.

This model can be also interpreted using network theory, adding and removing nodes from the graph in function of the birth and death rates. In Fig6 we have integrated, using the Gillespie algorithm, the model in a scale-free network (*Barabasi-Albert*) of 10^3 nodes. In Fig6**A** the system performs an epidemic outbreak as in his analogous case for the stochastic SIR model. The peak of the outbreak is larger and arrives faster, as we have observed before in the network SIR model. The topology of the network benefits the spreading. On the other hand we observe that the system does not reach the endemic equilibrium. This could be because the size the network is not large enough to reproduce the deterministic behaviour, and the fluctuations make the system to stabilize in the non-endemic equilibrium. Looking at the Fig6**B** we see that in this model still existing the threshold at which there is no epidemic outbreak.

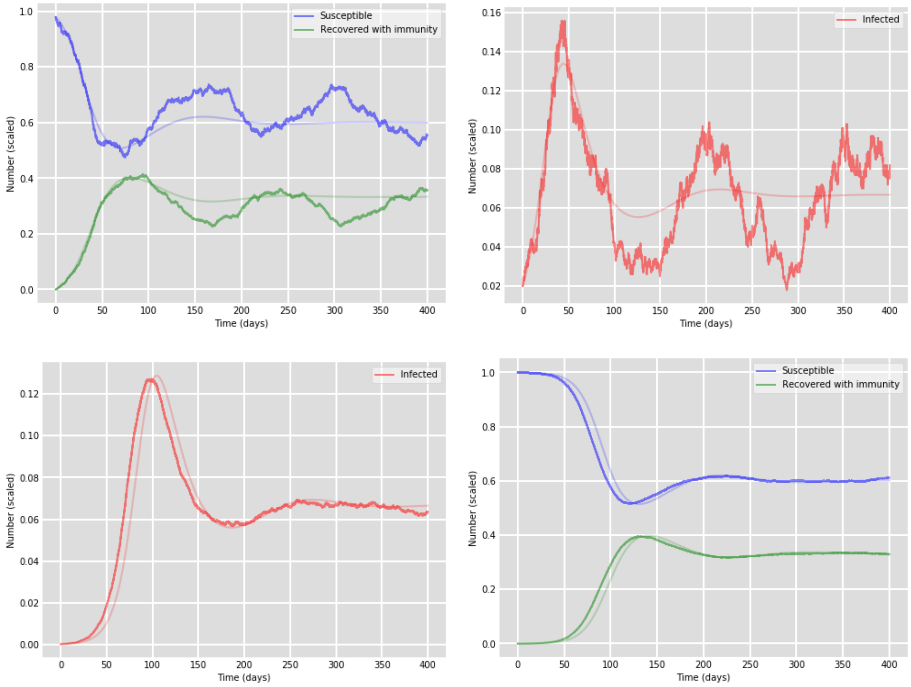

Figure 5: Integration of the demographic stochastic SIR model using a Gillespie algorithm, for $\beta = 0.2$ and a constant population. (Soft lines are the mean field solution).. **A:** Integration performed in a system with 10^3 individuals. **B:** Integration performed in a system with 10^5 individuals.

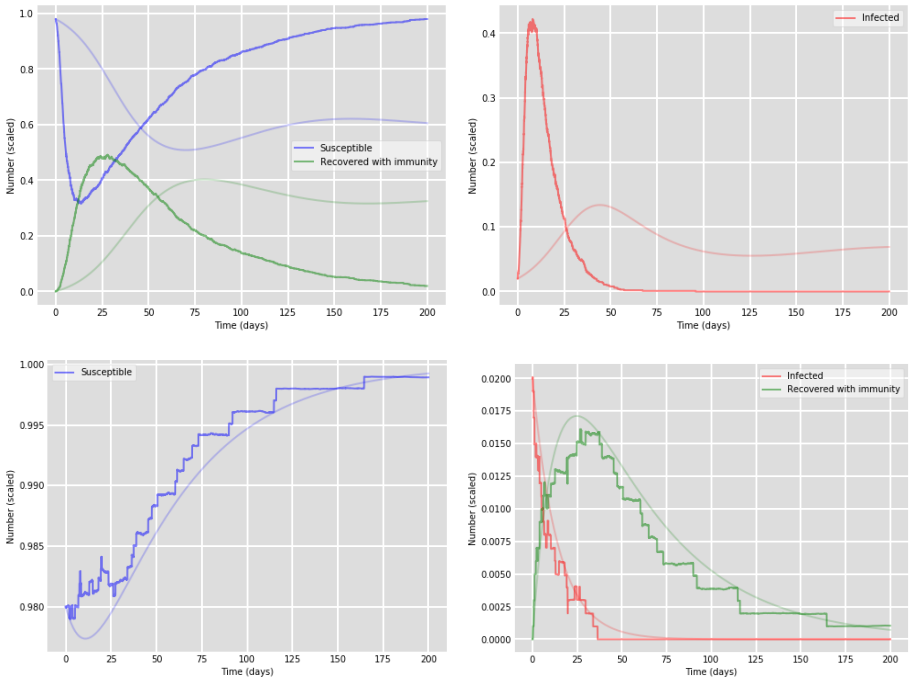

Figure 6: Integration of the demographic network SIR model in a *Barabasi-Albert* graph with 10^3 nodes, using a Gillespie algorithm, for a constant population ($\mu = \Delta$). **A:** For $\beta = 0.2$ there is epidemic outbreak. **B:** For $\beta = 0.05$ there is no epidemic outbreak.

6 Vaccination in a network SIR model

We can introduce to the model the effect of vaccinate the population in order to stop the epidemic outbreak. Once the spreading has started we could introduce a certain number of vaccines over the population to remove the susceptibles and cure the infecteds. We have implemented this mechanics in the network SIR model to study how the vaccinated population has to be chosen. We have vaccinated 5% of the population an 80% before that the epidemic outbreak of the network SIR model takes place. The results are really interesting. When the vaccinated nodes are chosen as those with highly degree in the network we achieve to reduce the impact of the infection (Fig7**A**), whereas if the vaccinated nodes are randomly selected, the solution is pretty similar to the original one (Fig7**B**). This is due to the topology of scale-free networks. We are working with highly clustered networks and removing the main hubs reduces significantly the effective spreading of the infection over the network.

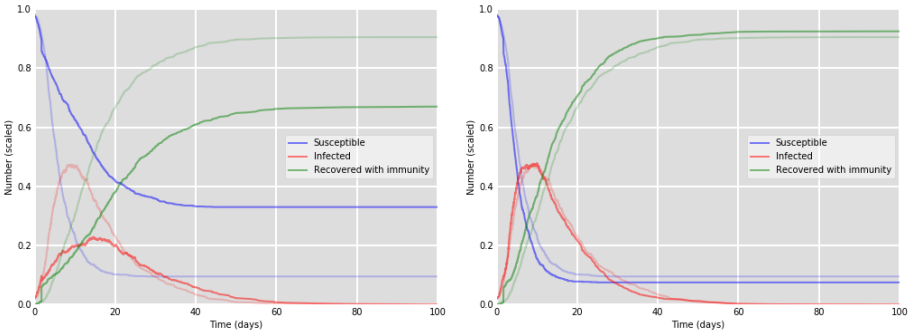

Figure 7: Effect of the vaccination on a *B-A model* of 10^3 nodes for $\beta = 0.2$. The population is maintained constant using $\mu = \Delta$. **Left:** 5% of the population is vaccinated, chosen among those with higher degree. **Righ:** 5% of the population is vaccinated, choosing randomly among all nodes.

7 References

[1] *Stochastic simulation of epidemics* Level 2 module in "Modelling course in population and evolutionary biology" - (701-1418-00). Roland Regoes, Institute of Integrative Biology ETH Zurich.
[2] Keeling MJ, Eames KT. Networks and epidemic models. Journal of the Royal Society Interface. 2005;2(4):295-307. doi:10.1098/rsif.2005.0051.
[3] Bastin, G. (2012). Lectures on Mathematical Modelling of Biological Systems.
[4] *Notes of the Master of Physics of Complex Systems course: Stochastic Processes.* Raúl Toral, November 19, 2015.
[5] Tao, Z., Zhongqian, F., Binghong, W. (2006). *Epidemic dynamics on complex networks.* Progress in Natural Science, 16(5), 452-457.
[6] Pastor-Satorras, R., Vespignani, A. (2001). *Epidemic spreading in scale-free networks.* Physical review letters, 86(14), 3200.
[7] *Spatial Stochastic SIR Model for Spread of Infectious Diseases.* John Lang, April 2012.